Article

Sugarcane Stem Node Recognition in Field by Deep Learning Combining Data Expansion

Wen Chen ¹, Chengwei Ju ², Yanzhou Li ¹, Shanshan Hu ^{1,*} and Xi Qiao ^{1,3,*}

¹ College of Mechanical Engineering, Guangxi University, Nanning 530004, China; CW934048393@outlook.com (W.C.); lyz197916@126.com (Y.L.)

² Guangxi Special Equipment Supervision and Research Institute, Nanning 530200, China; juchengwei1977@163.com

³ Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China

* Correspondence: chloe_2000@163.com or hsswhh@gxu.edu.cn (S.H.); qiaoxi@caas.cn (X.Q.)

Abstract: The rapid and accurate identification of sugarcane stem nodes in the complex natural environment is essential for the development of intelligent sugarcane harvesters. However, traditional sugarcane stem node recognition has been mainly based on image processing and recognition technology, where the recognition accuracy is low in a complex natural environment. In this paper, an object detection algorithm based on deep learning was proposed for sugarcane stem node recognition in a complex natural environment, and the robustness and generalisation ability of the algorithm were improved by the dataset expansion method to simulate different illumination conditions. The impact of the data expansion and lighting condition in different time periods on the results of sugarcane stem nodes detection was discussed, and the superiority of YOLO v4, which performed best in the experiment, was verified by comparing it with four different deep learning algorithms, namely Faster R-CNN, SSD300, RetinaNet and YOLO v3. The comparison results showed that the AP (average precision) of the sugarcane stem nodes detected by YOLO v4 was 95.17%, which was higher than that of the other four algorithms (78.87%, 88.98%, 90.88% and 92.69%, respectively). Meanwhile, the detection speed of the YOLO v4 method was 69 f/s and exceeded the requirement of a real-time detection speed of 30 f/s. The research shows that it is a feasible method for real-time detection of sugarcane stem nodes in a complex natural environment. This research provides visual technical support for the development of intelligent sugarcane harvesters.

Keywords: sugarcane stem nodes; complex natural environment; YOLO v4; visual identification

Citation: Chen, W.; Ju, C.; Li, Y.; Hu, S.; Qiao, X. Sugarcane Stem Node Recognition in Field by Deep Learning Combining Data Expansion. *Appl. Sci.* **2021**, *11*, 8663. <https://doi.org/10.3390/app11188663>

Academic Editors:
Jesus Simal-Gandara, Jianbo Xiao and
Md Afjalus Siraj

Received: 25 July 2021

Accepted: 14 September 2021

Published: 17 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The area occupied by sugarcane planting in China ranks third in the world. However, the mechanisation of sugarcane harvesting is still at a relatively low level, for the main reasons that the mechanical harvesting destroys the stem nodes kept in the soil for the second year of growth, the impurity rate is high and the cutter is seriously worn by cutting into the soil. In contrast, although inefficient and labour-intensive, manual harvesting is widely adopted with good quality and flexibility. Therefore, it is necessary to improve the intelligence of sugarcane mechanical harvesting, and the recognition of the sugarcane cutting location is the first step toward intelligence.

Machine vision technology offers the possibility of identifying sugarcane stem nodes against a single background. Moshashai et al. [1] first studied the recognition of sugarcane stem nodes by comparing the diameter of different parts of the sugarcane and found that the diameter of the stem node was larger than the rest, which can be used to determine the position of the stem section. Shangping Lu et al. [2] proposed a feature extraction and recognition method of sugarcane stem nodes through the support vector machine

method by extracting features of the S and H component images in the HSV colour space of sugarcane segment pictures. However, the background of the sugarcane image was an ideal one with a single colour. The local mean [3] was another method used to identify the sugarcane stem node by filtering the image and image segmentation on H components of the HSV colour space, and it then found the maximum grey value as the stem node position. The experimental object was the image with only a single sugarcane stem node. Weizheng Zhang et al. [4] studied the method of identifying and locating sugarcane stem nodes based on high spectral light imaging technology. Its recognition range was limited to the area around the sugarcane stem node, and the recognition accuracy was 98.33%. Yanmei Meng et al. [5] proposed a sugarcane node recognition algorithm based on multi-threshold and multi-scale wavelet transform, even though the sugarcane could only be identified by stripping the sugarcane leaves in advance to expose the sugarcane node. Deqiang Zhou et al. [6] proposed a method of sugarcane stem node recognition based on Sobel edge detection to satisfy the working requirements of the sugarcane seed cutter. Jiqing Chen et al. [7] proposed a sugarcane nodes identification algorithm based on the sum of local pixels of the minimum points of vertical projection function to analyse the recognition of a single node and double nodes.

The methods mentioned above mainly relied on the traditional image-processing machine version method. It is clear that the algorithm of the machine version can identify the object mainly by analysis of the reflecting light and perspective light on the surface of the object, which needs to work in a simple environment. Although some of them cannot meet the requirements of real-time detection against a complex background or deal with sugarcane stem node recognition amidst sugarcane leaf wrapping, the research still puts forward feasible visual identification techniques for sugarcane seeding or harvesting machines.

Unlike traditional image processing that focuses on image feature recognition, deep learning is a learning method driven by big data, which has been widely used in agriculture [8–10], crop classification [11], crop image segmentation [12] and crop object detection [13]. Parvathi, S et al. [14] used the Faster R-CNN based on a deep learning algorithm to detect coconut in a complex background. Liang, Cuixiao et al. [15] studied the performance of the SSD algorithm in identifying lychee fruits and litchee branches at night. In 2020, Biffi, Leonardo Josué Mitishita Edson et al. [16] studied the detection of apples in apple orchards based on the RetinaNet algorithm through the ground remote sensing system. Wu, Dihua et al. [17] discussed and compared the identification accuracy of the apple flower in the field based on the You Only Look Once v4(YOLO v4) algorithm and YOLO v3 algorithm. Deep learning is a new and efficient method for intelligent cultivation in sugarcane planting. In 2020, J. Scott [18] used deep learning to study the furrow mapping of sugarcane billet density, and Srivastava, S [19] proposed an approach based on deep learning for sugarcane disease detection. These research studies demonstrated that deep learning has a stronger identification ability in the natural environment and sugarcane field. In 2019, Shangping Li et al. [20] introduced the object detection technology for sugarcane stem node recognition based on deep learning, which was applied in the sugarcane cutting process for the first time. It used an improved YOLO v3 network to establish an intelligent recognition convolutional neural network model. However, the sugarcane samples were pre-processed by removing the leaves manually first in a pre-processed single-colour background environment. Table 1 shows the relevant studies by the above-mentioned scholars on sugarcane stem node recognition. Although deep learning has been adopted in other crop recognition applications, it is still rarely used in sugarcane stem node recognition.

The visual identification of sugarcane stem nodes in the complex natural environment still goes unreported due to the following difficulties: (1) The complex lighting conditions in the natural environment and unstable sunlight during the day reduce the image quality and affect the accuracy of the detection algorithm; (2) sugarcanes grow in clusters, and some of the sugarcane stem nodes are more or less covered by leaves; and (3) the diversity of the biological characteristics of sugarcane, including different stalk diameters and peel colour, increase the difficulty of identification. In order to solve these problems, this paper

proposes a sugarcane stem node recognition algorithm based on deep learning driven by big data in the natural environment. The big data acquisition experiments were conducted at a real sugarcane farm, and the big data samples of the sugarcane stem node consisted of different light conditions and different shooting angles using the data expansion technique and images lighting conversion. The object detection algorithm based on deep learning can learn and understand the characteristics of different sugarcane stem nodes in the natural environment by learning big data.

Table 1. Related research on sugarcane stem node recognition.

Author	Method	Precision (%)	Detection Time (s)	Remark
Shangping Lu et al. [2]	Classification and recognition method based on SVM	91.552	0.76	In a laboratory environment, sugarcane leaves were stripped off.
Yiqi Huang et al. [3]	Recognition method based on Radon transform	100	0.21	In a laboratory environment, sugarcane leaves were stripped off.
Weizheng Zhang et al. [4]	Hyperspectral imaging technology	98.31	/	In a laboratory environment, sugarcane leaves were stripped off.
Yanmei Meng et al. [5]	Recognition technology based on wavelet analysis	100	0.25	In a laboratory environment, sugarcane leaves were stripped off.
Shangping Li et al. [16]	An improved YOLO v3	90.38	0.0228	In a laboratory environment, sugarcane leaves were stripped off.
Deqiang Zhou et al. [6]	Sobel edge detection method	93	0.539	In a laboratory environment, sugarcane leaves were stripped off.
Jiqing Chen et al. [7]	Vertical projection function method	98.5	0.21	In a laboratory environment, sugarcane leaves were stripped off.
This paper	YOLO v4	95.12	0.0145	In a natural environment, with sugarcane leaves not stripped

The rest of this article is organised as follows. The second section introduces the experimental procedure and data processing, including image acquisition, data expansion and the creation of image datasets. The third section introduces the sugarcane stem node detection model based on the YOLO v4 [21] algorithm in the natural environment. This algorithm is currently the best one-stage detection algorithm. The fourth part is the experimental part, which mainly discusses and analyses the experimental results. The last is the conclusion and prospect of this article.

2. Image Data Acquisition and Processing

Figure 1 shows the systematic research route of this study.

2.1. Image Data Acquisition

The images of the bottom of the sugarcane were collected from the sugarcane farm in Fusui County, Guangxi, China. The sugarcane variety was Guitang No. 49, the sugarcane was in the mature stage and the average stem diameter was about 2.5 cm. The sugarcane was grown in the open air and planted side by side according to the requirements for mechanical harvesting. In order to match the diversity of the sample environment, images were collected at 8:00, 12:00 and 18:00, and the lighting conditions include side light, forward light and back light. These were the three moments when the light intensity changed the most in the daytime. During this period, the sugarcane photos in different light directions can be obtained by adjusting the camera shooting angle. During image acquisition, the shooting direction of the camera can simulate the forward light, side light and back light by setting in the same, vertical and opposite directions as the light propagation direction. Considering that the camera's shooting angle will affect the detection performance, images were collected from multiple shooting angles during the image acquisition process.

The image set collected was composed of images of one single sugarcane stem node and images of multiple sugarcane stem nodes at a ratio of 1:3 to improve the robustness of the algorithm model. These two types of images are shown in Figure 2. Then, 1600 images were expanded to 8000 images using data expansion to generate the training data set and testing data set. The training data set was 7200 images, and the testing dataset was 800 images at the ratio of 9:1.

Figure 1. The systematic research route.

Figure 2. Collected image dataset.

2.2. Image Data Expansion

Because the angle and intensity of the light change greatly during the day, the ability of the neural network to process the images collected at different times of the day depended on the integrity of the training dataset. In order to enhance the diversity of the data and improve the recognition ability of the model under different images, the collected images were pre-processed with a random colour, brightness, rotation and a mirror flip. In this experiment, the programming language Python3.6 was used to implement the data expansion, the framework was PyCharm and the libraries used were Pillow, Numpy and OpenCV. The images processed are shown in Figure 3.

Figure 3. Image expansion methods.

2.2.1. Data Expansion by the Random Colour Method

Human beings recognise objects through the visual system, which is not affected by the changes of light and colour on the surface of objects, but a visual imaging device does not have such ability. Different lighting conditions will cause a certain deviation between the image colour and the true colour. Random colour processing of the image can further eliminate the influence of ambient light and improve the robustness of the detection model. The colour of the images was randomly adjusted by changing the saturation, sharpness, contrast and brightness of the image, and superimposing it to achieve the effect of random colour processing.

2.2.2. Data Expansion by the Image Rotation and Flip Method

In order to further extend the image dataset, the original image was rotated 30 degrees and flipped. Table 2 shows the number of images in the dataset after rotation and flip.

Table 2. Number of images generated by data expansion.

	Raw Data	Flip	Random Colour	Rotation	Brightness	Total
Morning images	400	400	400	400	400	2000
Noon images	800	800	800	800	800	4000
Evening images	400	400	400	400	400	2000

2.2.3. Data Expansion by the Image Brightness Method

In sugarcane fields in a wild environment the sugarcane leaves often block out the sun, which will result in insufficient light at the bottom of the sugarcane. Using the method of image brightness enhancement to expand the dataset, it is possible to simulate the condition of making up for the lack of illumination using an added artificial light. These extended datasets can compensate for the small variation of illumination intensity due to the short collection time. The number of images processed is shown in Table 2.

2.3. Image Annotation and Data Set Generation

The sugarcane images were manually labelled with LabelImg with bounding boxes drawn, classified into categories and saved in PASCAL VOC format. Marked rectangles were used to identify the sugarcane stem nodes. Data with insufficient or unclear pixel areas were not used to prevent overfitting in the neural network. The complete dataset is shown in Table 2.

3. Methodology

3.1. YOLO v4

The YOLO network is a one-stage object detection algorithm of a deep learning method that converts the detection problem into a regression problem. Compared with the Faster Region-based Convolutional Neural Network (Faster-RCNN) [22], it does not need a region proposal network, and can directly generate bounding box coordinates and the probability of each category through regression. This end-to-end object detection greatly improves the detection speed. YOLO v4 is the latest algorithm of the YOLO series and is regarded as an improved version of YOLO v3 [23]. Compared with YOLO v3, it adopts Mosaic data expansion in data processing, and optimises the backbone, network training, activation function and loss function, which is faster than YOLO v3, and achieves the best balance between accuracy and speed in these real-time object detection algorithms.

As shown in Figure 4, the YOLO v4 network uses the open-source neural network framework Centre and Scale Prediction Darknet53 (CSPDarknet53) [24] as the main backbone network for training and extracting image features; then, the Path Aggregation Network (PANet) [25] was used as the neck network to better integrate the extracted features; the head was the same as YOLO v3's method of detecting objects. The main modules of the sugarcane stem node detection model based on YOLO v4 in the complex natural environment were as follows:

- (1) Convolution, Batch Normalisation and Mish (CBM) is a convolutional layer, batch normalisation layer and Mish activation function. This module replaced the activation function in Convolution, Batch normalisation and Leaky-ReLU (CBL) of YOLO v3 with Mish, which is the most common module in YOLO v4.
- (2) CSPDarknet53 is the backbone of YOLO v4, which mainly consists of CBM and Centre and Scale Prediction (CSP). CSP is composed of CBM and the Res module, while the Res module is mainly composed of two CBM modules. Figure 4 shows their specific structure. It can enhance CNN's learning ability by dividing low-level features into two parts and then fusing cross-level features.
- (3) Spatial Pyramid Pooling (SPP) [26] is a spatial pyramid pooling layer, which mainly converts convolutional features of different sizes into pooled features of the same length.
- (4) PANet strengthens the entire feature hierarchy through the bottom-up path and uses accurate bottom-level positioning signals to shorten the information path between the bottom-level and top-level features.

3.2. YOLO v4 Algorithm Training Process

In order to realise the rapid detection of sugarcane stem nodes based on YOLO v4, the model weights of YOLO v4 were pre-trained on the Microsoft Common Objects (Context MS CoCo) dataset, and the model parameters of the network input size, number of categories, batch size and learning rate were fine-tuned. The total number of training epochs was 84, and the first 25 epochs were freezing training, which can ensure that the initial weight was not destroyed and speeds up the training. The main settings are shown in Table 3.

Table 3. Parameter settings.

Parameter	Value
Input size	416×416
Learning rate (freeze epoch)	1×10^{-3}
Learning rate	1×10^{-4}
Batch size (freeze epoch)	8
Batch size	2
Classes	1

Figure 4. Detection of sugarcane stem nodes in a complex natural environment based on YOLO v4.

The training set and testing set were used to train and test the YOLO v4 sugarcane stem node detection model. As shown in Formulas (1)–(4), the loss function used to train the YOLO v4 sugarcane stem node detection model mainly included the position loss of the bounding box, the confidence loss and the classification loss.

$$Loss = L_{CIoU} + L_{confidence} + L_{class} \quad (1)$$

$$L_{CIoU} = 1 - IOU + \frac{d^2}{c^2} + \alpha v \quad (2)$$

The c and d in Formula (2) are the distance between the centres of the two bounding boxes and the diagonal distance of their union, respectively.

$$L_{confidence} = \sum_{i=0}^{S^2} \sum_{j=0}^B K[-\log(p) + BCE(\hat{n}, n)] \quad (3)$$

$$L_{class} = \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{noobj} [-\log(1 - p_c)] \quad (4)$$

S is the number of grids and B is the anchor number corresponding to each grid in Formulas (3) and (4).

$$IOU(A, A') = \frac{A \cap A'}{A \cup A'} \quad (5)$$

where IOU , as an abbreviation for Intersection over Union, is the ratio of the intersection and union of the ground truth (A) and the predicted value boundary boxes (A') in Formula (5).

$$BCE(\hat{n}, n) = -n \log(n) - (1 - \hat{n}) \log(1 - n) \quad (6)$$

where BCE represents the cross-entropy loss function of the true value (n) and the predicted value ($\hat{n}$).

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (7)$$

where w^{gt} and h^{gt} are the height and width of the ground truth bounding box, and w and h represent the width and height of the predicted bounding box.

Formula 8 was derived jointly from Formulas (5) and (7).

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (8)$$

$$K = 1_{i,j}^{obj} \quad (9)$$

In Formula (9), K stands for weight. If there is an object in K , its value is 1. p is the probability that the detection object is a sugarcane stem node.

Figure 5 shows the total loss function curve during training. In the initial training stage of the sugarcane stem node detection model, the model learning efficiency was high, and the training converged fast. With the deepening of the training, the slope of the training curve decreased gradually. When the number of training iterations reached 80, the model learning gradually stabilised.

Figure 5. Total loss function curve of model training.

The detection result of sugarcane stem nodes based on YOLO v4 is shown in Figure 6. In addition to the images processed by random colours, the algorithm can detect the cane stem node from the original image and three kinds of data-enhanced images, which proved that the algorithm had a high accuracy. In Figure 6e, the lowest sugarcane stem node was over-exposed after random colour processing, so it could not be identified.

Figure 6. Detection results of 4 kinds of data-expanded images based on YOLO v4.

3.3. Performance Evaluation Index of Algorithm Model

Five commonly used indicators, precision P, recall rate R, mAP (Formula (14)), detection speed and F1 (Formula (12)), were used to verify the performance of the model. For a binary classification problem, the samples can be divided into four types according to the combination of the true category and the predicted category of the sample: TP (True Positive), FP (False Positive), TN (True Negative) and FN (False Negative). In this paper, when $\text{IOU} \geq 0.5$, it was a True Positive; when $0 < \text{IOU} < 0.5$, it was a False Positive. When $\text{IOU} = 0$, the background was detected, and it was regarded as a True Negative. When $\text{IOU} \geq 0.5$, and the prediction result considers that the IOU was less than 0.5, it was a False Negative. The confusion matrix for the classification results is shown in Table 4.

Table 4. Confusion matrix for the classification results.

Labelled	Predicted	Confusion Matrix
Positive	Positive	TP
Positive	Negative	FN
Negative	Positive	FP
Negative	Negative	TN

The precision P and recall rate R were defined as Formulas (10) and (11). P was used to describe the proportion of the samples where the prediction is positive that the prediction is positive. R was used to describe the proportion where the labelling was positive that the prediction was positive. The higher the values of these two, the better the performance

of the algorithm. Using the precision P as the vertical axis and the recall rate R as the horizontal axis, the precision recall (PR) curve was achieved.

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

The F1 score is a reference value derived from Recall and Precision, and its value is usually close to the smaller of the two. If the F1 score is high, it indicates that both the Recall and Precision are high, so it was hoped to obtain a high F1 score. The F1 score is defined as:

$$F1 = \frac{2 \times P \times R}{P + R} \quad (12)$$

The average precision (AP) can show the overall performance of a model under different score thresholds. In this paper, AP was obtained by averaging the precision value on the PR curve, and was defined as Formula (13). mAP was the sum of AP values for all categories/the number of categories, and C was the number of categories. Since only sugarcane stem nodes were detected in this paper, $C = 1$ was used in this study.

$$AP = \int_0^1 P(r) dr \quad (13)$$

$$mAP = \frac{\sum_{i=1}^C AP(C)}{C} \quad (14)$$

The detection speed (f/s) is the derivative of the computational time required by each method to recognise one sample. All the aforementioned methods were coded and developed in Python 3.6, and the deep learning framework was Keras. A workstation with a 2.3 GHz Intel 5218 $\times$ 2 processor, 64 GB RAM and 11GB NVIDIA RTX 2080Ti GPU was used for calculation and images processing.

4. Result and Discussion

4.1. The Recognition Effect of Different Algorithms

Four object detection algorithms, Faster R-CNN, SSD300 [27], RetinaNet [28] and YOLO v3, were selected and compared with the YOLO v4 algorithm to verify the recognition effect of the algorithm. The backbones of these four algorithms were ResNet50, VGG16, ResNet50 and Darknet53.

The training set was applied to the above five algorithms, and the test set was employed to evaluate the performance of the different detection algorithms. The P-R curves of the different algorithms are shown in Figure 7, which is a two-dimensional curve with precision and recall as the vertical and horizontal coordinates. When the P-R curve of one algorithm was surrounded by another algorithm, the latter performed better than the former. In Figure 7, except for the YOLO v3 algorithm and YOLO V4 algorithm, the curves of the other three algorithms all approached the coordinate point (1,0) at the end. Combined with Formulas (10) and (11), it was clear that the false detection rates (FP) of the Faster-RCNN, SSD300 and RetinaNet algorithms on the test set were all fairly high.

The F1 scores varying with the confidence threshold score are shown in Figure 8. The F1 scores are a harmonic average between the results of precision and recall, and range from 0 to 1, where 1 represents the best output of the model and 0 represents the worst output of the model. When the threshold value was set as 0.5 in this paper, the F1 value of YOLO v3 and YOLO v4 achieved the highest scores, indicating that the optimal output of the algorithm can be obtained when the two methods simultaneously meet the requirements of high precision and high recall rates.

Figure 7. The P-R curves of different algorithms.

Figure 8. The F1 score curves of different algorithms.

Table 5 was the statistical results of different algorithms. The results showed that the AP of these five object detection algorithms were 78.87%, 88.98%, 90.88%, 92.69% and 95.17%; the detection speeds were 11 f/s, 62.5 f/s, 40.18 f/s, 72 f/s and 69 f/s, respectively.

Table 5. Statistical results of different algorithms.

Algorithm	Precision (%)	Recall (%)	AP (%)	Detection Speed (f/s)	F1
Faster R-CNN	85.46	85.60	78.87	11	0.69
SSD300	93.17	81.20	88.98	62.5	0.87
RetinaNet	92.34	85.33	90.88	40.18	0.89
YOLO v3	94.40	85.31	92.69	72	0.90
YOLO v4	95.12	84.90	95.17	69	0.90

In terms of detection speed, although YOLO v3 was slightly faster than YOLO v4, both of them far exceeded the real-time detection requirement of 30 f/s. As for the AP results, YOLO v4's AP is 16.3%, 6.19%, 4.29% and 2.48% higher than the other four algorithms, respectively. Through the analysis of the test results, it can be seen that the detection accuracy of YOLO v4 for sugarcane stem nodes was higher than the other four algorithms, and the detection accuracy was very close to the fastest algorithm. It was clearly more in line with the requirements of sugarcane stem node recognition in the complex natural environment.

4.2. Comparative Experiments of Recognition under Different Lighting Factors

The light environment will change during the continuous harvesting of sugarcane. In this experiment, different shooting time periods (morning, noon and nightfall) were used as control variables to represent different illuminance levels, which were respectively oblique strong light, direct strong light and oblique weak light. The number of images in each time period was 100. The statistical detection results are shown in Table 6, and some of the image detection results are shown in Figure 9.

Table 6. Statistical detection results of five algorithms on stem nodes of sugarcane at different times.

Times	Algorithm	Precision (%)	Recall (%)	AP (%)	F1
Oblique strong light	Faster R-CNN	55.12	81.71	70.83	0.66
	SSD300	94.35	84.44	90.62	0.89
	RetinaNet	90.46	84.82	88.33	0.88
	YOLO v3	94.96	87.94	93.36	0.91
	YOLO v4	97.32	89.34	95.70	0.93
Direct strong light	Faster R-CNN	59.38	87.14	81.24	0.71
	SSD300	95.99	81.25	93.90	0.88
	RetinaNet	93.04	86.32	91.86	0.90
	YOLO v3	94.67	86.74	93.15	0.91
	YOLO v4	96.93	84.46	95.41	0.90
Oblique weak light	Faster R-CNN	58.56	86.41	79.88	0.70
	SSD300	95.98	78.10	92.76	0.86
	RetinaNet	92.44	85.61	91.12	0.89
	YOLO v3	94.51	85.91	92.94	0.90
	YOLO v4	97.74	84.64	95.04	0.91
Average	Faster R-CNN	57.69	85.09	77.32	0.69
	SSD300	95.44	81.26	92.43	0.877
	RetinaNet	91.98	85.58	90.44	0.89
	YOLO v3	95.05	86.86	93.15	0.907
	YOLO v4	97.33	86.15	95.38	0.913

It can be seen from Table 6 that the precision, AP and F1 score of the YOLO v4 algorithm were the highest. The intensity of illuminance had a great influence on the accuracy of all the algorithms. The key factor determining the accuracy of the algorithm was whether the stem nodes were under direct strong light. The detection accuracy was reduced when the light was oblique strong light or oblique weak light. Therefore, it is recommended that the intelligent sugarcane harvester can increase its illumination device to improve the detection accuracy when working continuously in dim daytime light.

It can be seen from Figure 9 that the colour and texture of the sugarcane stem nodes were clear and easy to recognise in the morning and at noon. At nightfall, due to the dimming of the illuminance and the shade from the branches and leaves, the illuminance of the sugarcane peel was reduced greatly, although the object detection algorithm based on deep learning can still accurately identify the location of the sugarcane stem nodes.

4.3. The Recognition Effect of Different Data Expansion Methods

As mentioned above, four data expansion methods were used in this article: Rotation, mirror flip, random colour processing and brightness enhancement. In order to verify the effect of these four data expansion methods on the performance of the algorithm, the variable control method was applied to delete the image data corresponding to each data expansion method in the training set, then the testing set was used to test the trained algorithm model and the YOLO v4 algorithm detection results were obtained. The results are shown in Table 7.

Figure 9. Image detection results of five algorithms on sugarcane stem nodes in different time periods.

Table 7. The recognition effect of different data expansion methods.

Data Expansion Methods	Precision (%)	Recall (%)	AP (%)	F1
Dataset after expansion	95.12	84.90	95.17	0.9
Remove rotation	91.41	69.80	78.48	0.79
Remove mirror flip	91.65	83.51	89.44	0.87
Remove random colour processing	90.78	74.33	85.90	0.82
Remove brightness enhancement	90.65	81.59	85.93	0.86

From Table 7, the method of rotation was very helpful to improve the detection accuracy. By deleting the images produced by the rotation method, the AP of the YOLO v4 detection model was reduced by 16.69%, and the F1 score was reduced by 0.11.

The method of flipping had the least impact on improving the detection accuracy. After removing the mirror flip images, the performance of the training model was only slightly lower than that of the complete dataset. The AP of the YOLO v4 detection model was reduced by 5.73%, and the F1 score was reduced by 0.03.

Compared to the dataset without the images processed by a random colour, the model trained with the complete dataset had higher detection accuracy. After the removal of the random colour processing from the training set, the AP of the YOLO v4 detection model decreased by 9.27% and the F1 score decreased by 0.02. This indicated that random colour processing was very beneficial to improve the robustness of the model.

The recognition model without images processed by brightness enhancement was worse than that of the model trained with the complete dataset. The AP of the YOLO v4 detection model was reduced by 9.27% and the F1 score was reduced by 0.04. Bright-

ness enhancement helped the model adapt to the lighting conditions of the complex natural environment.

4.4. Comparison with Previous Related Recognition Methods

In 2014, Girshick et al. proposed the RCNN (Region-based Convolutional Neural Network) [29] algorithm, which opened up a new era of object detection algorithms based on deep learning, and deep learning technology had also begun to be applied to the agricultural field on a large scale [8]. The previous methods of intelligent identification of sugarcane stem nodes have been fully discussed in Table 1 above, but they did not address the impact of lighting changes, sugarcane leaves and biological characteristics on recognition in complex environments. The research of this paper focuses on the recognition of sugarcane stem nodes in the field under the complex natural environment, which is still not understood at present. In order to improve the robustness and detection accuracy of the algorithm model, the data expansion method was used to enrich the datasets and simulate sugarcane images under different light conditions. Table 1 shows our research results on the recognition of sugarcane stem nodes.

In Table 1, comparing this paper with previous studies, we can find that deep learning technology has the advantage of not only recognising image features but also understanding the image content; this technology can detect sugarcane stem nodes more than 10 times faster than machine vision technology [3], which can thus satisfy the requirements of real-time detection.

At the same time, it is worth noting that, after fully considering the influence of light conditions, sugarcane leaves and biological characteristics in the complex environment, the detection speed of this paper's method was twice as fast as that of the YOLO v3 method on a simple background [20], and the accuracy was 4.74% higher than it too [20].

5. Conclusions

The object detection algorithm for sugarcane stem node recognition based on YOLO v4 in the natural environment was introduced in this paper for the first time and achieved rapid and accurate recognition of sugarcane stem nodes during harvest in the natural environment, while the robustness and generalisation ability of the algorithm were improved by the dataset expansion method to simulate different illumination conditions. The images were collected in different lighting conditions of side light, forward light and back light. The impact of the data expansion and lighting conditions at different times of the day on the detection results of sugarcane stem nodes was discussed, and the superiority of YOLO v4, which performed best in the experiment, was verified by comparison with four different deep learning algorithms, namely Faster R-CNN, SSD300, RetinaNet and YOLO v3. The main conclusions are as follows.

In the absence of a large amount of data, a data expansion method was adopted by simulating different illumination conditions and different shooting angles to train the detection model of sugarcane stem node recognition based on YOLO v4. The 1600 original images were expanded to 8000 images using data expansion to generate the training dataset and testing dataset. Through this method, the robustness of the model was effectively improved.

The AP of the object detection algorithm based on YOLO v4 was the highest, at 95.17%. Although the detection speed of YOLO v3 (72 f/s) was slightly faster than that of YOLO v4 (69 f/s), both of them far exceeded the real-time detection requirement of 30 f/s.

By comparison with the previous studies on sugarcane stem node recognition, the object detection algorithm based on YOLO v4 in a complex natural environment can detect sugarcane stem nodes wrapped by leaves more than 10 times faster than machine vision technology in a pre-processed single-colour background environment. Meanwhile, after fully considering the influence of light conditions, sugarcane leaves and biological characteristics in a complex environment, the detection speed of this paper's method was twice as fast as the previous method using YOLO v3 in a pre-processed single-colour

background, and the accuracy was 4.74% higher too. The result indicated that the detection method based on YOLO v4 was feasible for fast and accurate detection of sugarcane stem nodes in the complex natural environment. This method provides effective visual technical support for the intelligent sugarcane harvester.

Author Contributions: Conceptualisation, W.C.; methodology, W.C.; software, W.C.; validation, S.H., C.J., Y.L. and X.Q.; formal analysis, W.C.; investigation, W.C. and S.H.; resources, S.H.; data curation, W.C.; writing—original draft preparation, W.C.; writing—review and editing, S.H., Y.L. and X.Q.; visualisation, W.C.; supervision, S.H., C.J. and X.Q.; project administration, S.H. and C.J.; funding acquisition, S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Nature Science Foundation of China (No. 51965004 & 51565005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The author would like to thank the financial support provided by the National Nature Science Foundation of China, and Chengwei Ju, Yanzhou LI and Xi Qiao for writing advice. Most of all, Wen Chen wants to thank the first corresponding author Shanshan Hu for the constant support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Moshashai, K.; Almasi, M.; Minaei, S.; Borghei, A.M. Identification of sugarcane nodes using image processing and machine vision technology. *Int. J. Agric. Res.* **2008**, *3*, 357–364. [\[CrossRef\]](#)
2. Lu, S.; Wen, Y.; Ge, W. Recognition and features extraction of sugarcane nodes based on machine vision. *Trans. Chin. Soc. Agric. Mach.* **2010**, *41*, 190–194. [\[CrossRef\]](#)
3. Huang, Y.; Qiao, X.; Tang, S.; Luo, Z.; Zhang, P. Location and experiment of characteristic distribution of sugarcane stem nodes based on Matlab. *Trans. Chin. Soc. Agric. Mach.* **2013**, *44*, 93–97, 232. [\[CrossRef\]](#)
4. Zhang, W.; Zhang, W.; Zhang, H.; Chen, Q.; Ding, C. Research on identification and location method of sugarcane node based on hyperspectral imaging technology. *J. Light Ind.* **2017**, *32*, 95–102. [\[CrossRef\]](#)
5. Meng, Y.; Ye, C.; Yu, S.; Qin, J.; Zhang, J.; Shen, D. Sugarcane node recognition technology based on wavelet analysis. *Comput. Electron. Agric.* **2019**, *158*, 68–78. [\[CrossRef\]](#)
6. Zhou, D.; Fan, Y.; Deng, G.; He, F.; Wang, M. A new design of sugarcane seed cutting systems based on machine vision. *Comput. Electron. Agric.* **2020**, *175*, 105611. [\[CrossRef\]](#)
7. Chen, J.; Wu, J.; Qiang, H.; Zhou, B.; Xu, G.; Wang, Z. Sugarcane nodes identification algorithm based on sum of local pixel of minimum points of vertical projection function. *Comput. Electron. Agric.* **2021**, *182*, 105994. [\[CrossRef\]](#)
8. Tian, H.; Wang, T.; Liu, Y.; Qiao, X.; Li, Y. Computer vision technology in agricultural automation—A review. *Inf. Process. Agric.* **2020**, *7*, 1–19. [\[CrossRef\]](#)
9. Anagnostis, A.; Tagarakis, A.; Kateris, D.; Moysiadis, V.; Sørensen, C.; Pearson, S.; Bochtis, D. Orchard Mapping with Deep Learning Semantic Segmentation. *Sensors* **2021**, *21*, 3813. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Anagnostis, A.; Tagarakis, A.; Asiminari, G.; Papageorgiou, E.; Kateris, D.; Moshou, D.; Bochtis, D. A deep learning approach for anthracnose infected trees classification in walnut orchards. *Comput. Electron. Agric.* **2021**, *182*, 105998. [\[CrossRef\]](#)
11. Arribas, J.I.; Sánchez-Ferrero, G.V.; Ruiz-Ruiz, G.; Gómez-Gil, J. Leaf classification in sunflower crops by computer vision and neural networks. *Comput. Electron. Agric.* **2011**, *78*, 9–18. [\[CrossRef\]](#)
12. Dias, P.A.; Tabb, A.; Medeiros, H. Apple flower detection using deep convolutional networks. *Comput. Ind.* **2018**, *99*, 17–28. [\[CrossRef\]](#)
13. Yamamoto, K.; Guo, W.; Yoshioka, Y.; Ninomiya, S. On Plant Detection of Intact Tomato Fruits Using Image Analysis and Machine Learning Methods. *Sensors* **2014**, *14*, 12191–12206. [\[CrossRef\]](#)
14. Parvathi, S.; Selvi, S.T. Detection of maturity stages of coconuts in complex background using Faster R-CNN model. *Biosyst. Eng.* **2021**, *202*, 119–132. [\[CrossRef\]](#)
15. Liang, C.; Xiong, J.; Zheng, Z.; Zhong, Z.; Li, Z.; Chen, S.; Yang, Z. A visual detection method for nighttime litchi fruits and fruiting stems. *Comput. Electron. Agric.* **2020**, *169*, 105192. [\[CrossRef\]](#)
16. Biffi, L.; Mitishita, E.; Liesenberg, V.; Santos, A.; Gonçalves, D.; Estrabis, N.; Silva, J.; Osco, L.P.; Ramos, A.; Centeno, J.; et al. ATSS Deep Learning-Based Approach to Detect Apple Fruits. *Remote Sens.* **2020**, *13*, 54. [\[CrossRef\]](#)

17. Wu, D.; Lv, S.; Jiang, M.; Song, H. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Comput. Electron. Agric.* **2020**, *178*, 105742. [[CrossRef](#)]
18. Scott, J.; Busch, A. Furrow Mapping of Sugarcane Billet Density Using Deep Learning and Object Detection. In Proceedings of the 2020 Digital Image Computing: Techniques and Applications (DICTA), Melbourne, Australia, 29 November–2 December 2020. [[CrossRef](#)]
19. Srivastava, S.; Kumar, P.; Mohd, N.; Singh, A.; Gill, F.S. A Novel Deep Learning Framework Approach for Sugarcane Disease Detection. *SN Comput. Sci.* **2020**, *1*, 1–7. [[CrossRef](#)]
20. Li, S.; Li, X.; Zhang, K.; Li, K.; Yuan, L.; Huang, Z. Improve the YOLOv3 network to improve the efficiency of real-time dynamic recognition of sugarcane stem nodes. *Trans. Chin. Soc. Agric. Eng.* **2019**, *35*, 185–191. [[CrossRef](#)]
21. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-J.M. YOLOv4 Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934. Available online: <https://arxiv.org/abs/2004.10934> (accessed on 1 September 2021).
22. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
23. Redmon, J.; Farhadi, A. YOLOv3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767. Available online: <https://arxiv.org/abs/1804.02767> (accessed on 1 September 2021).
24. Wang, C.-Y.; Liao, H.-Y.M.; Wu, Y.-H.; Chen, P.-Y.; Hsieh, J.-W.; Yeh, I.-H. CSPNet: A New Backbone that Can Enhance Learning Capability of CNN. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020; pp. 1571–1580.
25. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path Aggregation Network for Instance Segmentation. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018.
26. Saumitou-Laprade, P.; Vernet, P.; Vekemans, X.; Castric, V.; Barcaccia, G.; Khadari, B.; Baldoni, L. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916. [[CrossRef](#)]
27. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Volume 9905, pp. 21–37. [[CrossRef](#)]
28. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2999–3007.
29. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 24–27 June 2014; pp. 580–587. [[CrossRef](#)]